

Preprint del capítulo de libro:

Alcina, A., V. Soler y A. Estellés (2005): «Internet como instrumento para la documentación en terminología y traducción. Hacia las plataformas de recursos electrónicos para el traductor especializado», en D. Sales (ed.) *La Biblioteca de Babel: Documentarse para traducir*, Granada: Comares, 221-241. ISBN: 84-8151-032-7. <http://hdl.handle.net/10234/192441>
@ Las autoras

INTERNET COMO INSTRUMENTO PARA LA DOCUMENTACIÓN EN TERMINOLOGÍA Y TRADUCCIÓN. HACIA LAS PLATAFORMAS DE RECURSOS ELECTRÓNICOS PARA EL TRADUCTOR ESPECIALIZADO

M^a Amparo Alcina Caudet, Victoria Soler Puertes, Anna Estellés Palanca

Universitat Jaume I

1. Introducción¹

La traducción de un texto especializado exige documentarse previamente sobre esa especialidad para poder comprender el ámbito temático en el que se inscribe y acceder a su terminología, de forma que posteriormente se pueda expresar su contenido en la lengua de llegada. La tarea de la escritura del texto meta parece la más evidente entre las que se realizan durante la traducción, ya que de ella resulta directamente el producto, el texto traducido. Sin embargo, gran parte del trabajo de traducir reside principalmente en la comprensión del texto origen, y en esta tarea el dominio de conocimientos y destrezas de la documentación resultan imprescindibles.

Las dificultades para la completa comprensión del texto origen proceden de la complejidad del ámbito temático, la densidad y dificultad terminológica, pero también del registro utilizado por el autor del texto dependiendo tanto de su propio dominio del tema (experto o lego) como del receptor a quien dirige el texto, del tipo de texto (argumentativo, expositivo, descriptivo...) o de las convenciones estilísticas del género (un contrato, un manual de un aparato, un artículo científico, un prospecto de un medicamento, un documental, un informe técnico).

Tradicionalmente, las fuentes documentales que ha utilizado el traductor han sido las enciclopedias, los manuales y los diccionarios impresos, disponibles generalmente en bibliotecas especializadas. El objetivo principal era conseguir información concreta o general sobre un determinado proceso científico o tecnológico, una estrategia comercial o una nueva técnica curativa, por poner un ejemplo, así como conocer las equivalencias lingüísticas de los

¹ Este trabajo ha sido realizado gracias a la financiación del Ministerio de Ciencia y Tecnología (BFF2002-

términos que aparecían en el texto origen.

Actualmente, gracias al desarrollo de la red Internet, el traductor ha cambiado fundamentalmente dos aspectos en la fase de documentación. Por una parte, el cómo buscar información, y por otra parte, qué tipo de información puede buscar o encontrar. La búsqueda de información ha cambiado porque a) el traductor puede hacer las búsquedas sin salir de su despacho, si dispone de una conexión a Internet; y b) la información a la que puede acceder a través de Internet es notablemente superior en cantidad y variedad. Por otra parte, también ha cambiado respecto a qué tipo de información puede buscar porque ahora, junto a la búsqueda de contenidos temáticos y equivalencias terminológicas, el traductor puede obtener textos especializados de diferente tipo textual y niveles de especialización.

La gran cantidad y variedad de datos disponibles a través de Internet puede convertirse, sin embargo, en una trampa para el traductor. En primer lugar, hay que tener en cuenta que no todo lo que encontremos en Internet es información relevante ni fiable en su totalidad, lo que nos lleva a la necesidad de seleccionar y clasificar la información. En segundo lugar, la capacidad de asimilar información es limitada en el tiempo, por lo que se necesitan herramientas y técnicas que permitan extraer automáticamente la información precisa de grandes volúmenes de documentos rápidamente, de manera que el trabajo del traductor resulte eficaz.

Este nuevo panorama obliga al traductor a plantearse la fase de documentación de una manera diferente. Ahora la dificultad no está en hacer un gran acopio de datos, ya que eso ya nos lo ofrece Internet, sino cómo filtrar la información relevante, almacenarla masivamente y extraer de forma inequívoca los datos buscados. En este nuevo panorama, el conocimiento de instrumentos informáticos y la destreza en su manejo serán, junto a las técnicas documentales, compañeros inseparables del traductor especializado.

En este artículo presentamos los principales aspectos estratégicos de búsqueda de información en Internet y la ficha de evaluación que forman parte del protocolo diseñado en el marco de los proyectos CREC y TXTCeram. En estos proyectos se pretende, por una parte, formar al estudiante de traducción e interpretación en búsqueda y selección de información fiable y de calidad en Internet, y, por otra parte, la creación de corpus electrónicos de textos de especialidad en los ámbitos de cerámica y turismo.

Por último, hacemos una reflexión sobre el futuro de la documentación en Internet para el traductor especializado, y nos referimos a la necesidad de aunar los esfuerzos para

hacer más eficientes las tareas en la fase de información y documentación, proponiendo la creación de plataformas virtuales para la información del traductor especializado.

2. El proceso de documentación en Internet del traductor especializado

La principal ventaja que ofrece Internet a cualquier persona que busca información es el acceso inmediato a una gran cantidad de documentos a partir de la introducción de una palabra clave en un motor de búsqueda. Este procedimiento sin embargo puede resultar algo desesperante para el traductor especializado porque:

- a) Existen desigualdades importantes entre las informaciones disponibles en distintos idiomas. Mientras que es fácil encontrar documentos especializados de cualquier ámbito en inglés, no es tan fácil si la búsqueda la hacemos, por ejemplo, en catalán.
- b) En ocasiones, los contenidos encontrados no responden a una verdadera especialización, sino que son documentos de entretenimiento escritos por aficionados o principiantes y por tanto pueden ofrecer información de dudosa calidad bien sea por el contenido poco preciso o porque no se han actualizado debidamente.
- c) A menudo, un motor de búsqueda puede ofrecer gran cantidad de resultados, de los cuales no todos resultan afines al tema buscado a pesar de contener la palabra clave buscada.
- d) Por último, también puede darse el caso de que existiendo información relevante sobre el tema buscado, ésta no nos sea ofrecida por los motores de búsqueda. Hay que tener en cuenta, además, que no toda la información que se ofrece en Internet es gratuita, y por tanto, para acceder a cierta información será necesario estar suscrito a un determinado servicio.

Estos son algunos de los elementos que nos indican que no toda la información que encontraremos en Internet es fiable, y por tanto es necesario diseñar estrategias que aseguren al traductor la mayor cantidad de contenidos con el máximo nivel de calidad respecto a nuestros criterios de búsqueda (Bergeron y van Steenbergen, 2000). Por otra parte, también nos indica que a pesar de la gran cantidad y variedad de datos que podemos encontrar en Internet gratuitamente, hoy por hoy sigue siendo necesaria la información en formato impreso, ya que la información más relevante y especializada normalmente no está disponible en Internet de forma gratuita. Gran parte de las páginas de Internet disponibles gratuitamente contienen un contenido divulgativo. Si queremos acceder a contenidos más especializados sería necesario suscribirse a bases de datos, lo que normalmente por motivos de rentabilidad

económica está, hoy en día, fuera del alcance del traductor.

Por último, debemos tener en cuenta que las ventajas que nos ofrece hoy en día la informática permiten que los datos recopilados en formato electrónico puedan ser consultados y procesados de diferentes maneras para acceder más fácil y eficazmente a la información. De ahí que para el traductor sea necesario dividir el proceso de documentación en tres fases: 1) Búsqueda de información en Internet; 2) Evaluación de la información; y 3) Almacenamiento de la información.

2.1. La búsqueda de información en Internet

El traductor dispone de varias opciones para documentarse eficazmente a través de Internet. La forma más habitual de búsqueda de información es el uso de motores de búsqueda (como Google, Altavista, etc.), los listados de información previamente seleccionada por traductores y expertos, y la consulta a expertos a través de Internet.

2.1.1. Los motores de búsqueda y directorios

Al realizar una búsqueda a través de un motor, es conveniente dominar todos los aspectos que conciernen a la configuración y estructura de una página web, así como a las estrategias que utilizan estos motores. En la búsqueda a través de motores de búsqueda, el usuario introduce una o varias palabras clave relacionadas con aquello que busca y el sistema le devuelve un listado de resultados. La infraestructura del motor de búsqueda consiste básicamente en una base de datos donde se almacenan los principales datos de las páginas que son visitadas e indexadas automática y periódicamente por el motor para extraer toda esa información. Un directorio, en cambio, recopila la información de páginas web de forma organizada y que ha sido analizada generalmente por personas. La posibilidad de error en la clasificación de la información es menor en los directorios, pero la información que contienen es menos abundante. Conviene conocer bien las características de unos y otros cuando los utilizamos como fuentes de documentación (Budin y Wright, 2001).

En general, para el traductor siempre serán más útiles las búsquedas por palabras clave mediante los motores de búsqueda, más que los directorios (Bergeron y van Steenberg, 2000). El primer paso en la búsqueda es la elección del motor de búsqueda. Hay multitud de buscadores (el más conocido y utilizado actualmente es Google), pero no todos los motores de búsqueda son iguales. Difieren en la forma en que indexan la información y las opciones de búsqueda que ofrecen, el tamaño de la base de datos, la rapidez en la búsqueda, el diseño de la

interfaz, la forma en que ofrecen los resultados y la ayuda que ofrecen. Conviene consultar más de un motor de búsqueda, y conocer las distintas opciones de búsqueda que ofrecen. Entre las opciones, podremos encontrar la posibilidad de la inclusión o exclusión de palabras clave buscadas o no deseadas, la búsqueda sólo en páginas en una determinada lengua, el uso de comodines, la sensibilidad a mayúsculas o determinados diacríticos en las palabras clave, la búsqueda sólo en ciertos dominios o la exclusión de determinados dominios.

Tras seleccionar el motor de búsqueda más adecuado, el siguiente paso es formular la pregunta y seleccionar los resultados. Si los resultados no fueran los más adecuados, habrá que modificar la pregunta (Bergeron y van Steenberghe, 2000).

Habitualmente se utiliza la casilla de búsqueda del motor para indicar la palabra o palabras clave relacionadas con el tema del que buscamos información. Si queremos resultados más adecuados al perfil de sitios que queremos encontrar, podemos utilizar otras opciones del motor de búsqueda, como el uso de operadores booleanos o la restricción de la búsqueda según el idioma, el dominio al que queremos restringir nuestra búsqueda, etc.

Podemos utilizar una o más palabras clave en la búsqueda. Normalmente, si utilizamos más de una palabra clave, el motor mostrará todas las páginas que tengan una de las dos o más palabras clave. Si queremos que muestre sólo las que tengan las dos palabras clave, será necesario utilizar algún operador. Por otra parte, si lo que queremos es que obtenga casos en los que las dos palabras aparecen juntas y en el orden indicado, también tendremos que expresarlo mediante operadores determinados. Por último, también se puede hacer uso de comodines que pueden sustituir a una letra o varias.

Ejemplos:

- *cerámica industrial* Muestra páginas que tienen *cerámica*, *industrial* o ambas.
- *+cerámica + industrial* Muestra resultados que tienen las dos palabras
- *“cerámica industrial”* Muestra resultados que contienen la combinación *cerámica industrial*.
- *+cerámica –artística* Muestra resultados que contienen la palabra *cerámica* y no tienen la palabra *artística*.
- *baldos** Muestra resultados que contienen las palabras *baldosa*, *baldosas*, *baldosín*, *baldosines*.²

Según el buscador de que se trate, en lugar de estos operadores se pueden utilizar

² Los símbolos utilizados como comodines pueden ser distintos en cada buscador. Los más comunes son el

operadores booleanos *AND*, *OR*, *NOT*, cuya sintaxis también puede variar. Por eso, es importante leer atentamente los ficheros de ayuda.

Además de restringir las opciones en las palabras clave, también se puede restringir la búsqueda por otro tipo de criterios, como son los siguientes:

Cuando insertamos una palabra en la casilla de búsqueda de Google, el motor busca en el índice de palabras de todas las páginas indexadas en su base de datos. Para afinar la búsqueda, podemos solicitar que se limite a ciertos campos de la página (por ejemplo, su título), o que haga la búsqueda sólo en las páginas escritas en una determinada lengua, o alojadas con un determinado dominio (por ejemplo, .org). En la Figura 1, podemos ver cómo se ha configurado en Google una búsqueda avanzada con restricción de criterios, que se explican a continuación.

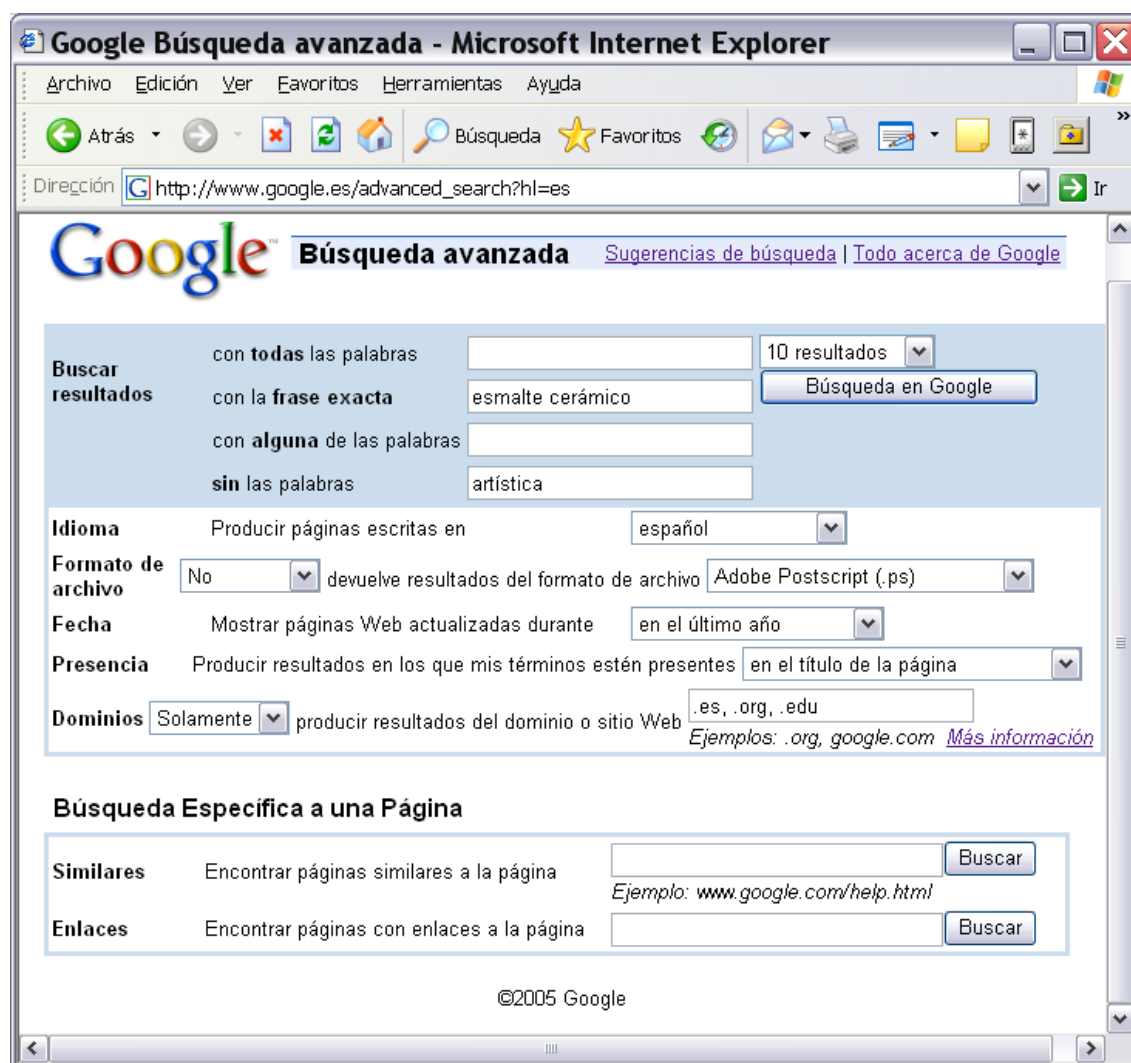


Figura 1. Ventana de búsqueda avanzada en Google

asterisco (*), el signo de interrogación (?) y el símbolo de tanto por ciento (%).

- A) Resultados. En cuanto a la palabra de búsqueda, se ha especificado la combinación “*esmalte cerámico*”, pero a la vez eliminándose aquellas páginas que contengan la palabra *artística*. Dado que la cerámica tiene diversas especialidades, como la cerámica industrial y la cerámica artística, al hacer la búsqueda en Internet resultan muchas páginas de ambas especialidades. Al incluir la palabra *artística* en las palabras no deseadas, intentamos eliminar aquellas páginas que se refieren a la especialidad artística.
- B) Idioma. En segundo lugar, se ha restringido la búsqueda a páginas que estén escritas en español, dado que es en esta lengua en la que necesitamos recopilar textos especializados.
- C) Formato de archivo. En cuanto al formato de archivo, se ha excluido la posibilidad de que se nos muestren ficheros en formato ps, es decir, postscript. El motor mostrará cualquier tipo de formato encontrado (html, htm, pdf, excel, etc.) pero no aquellos en postscript. Esta restricción puede ser útil por distintos motivos. En nuestro caso, por ejemplo, responde al hecho de la dificultad de tratar estos ficheros con las herramientas de tratamiento lingüístico disponibles. De este modo, eliminamos de antemano esos enlaces que no nos van a resultar útiles.
- D) Fecha. Google permite seleccionar contenidos que han sido actualizados en los últimos tres, seis o doce meses.
- E) Presencia. Podemos restringir la búsqueda de nuestras palabras clave a una determinada sección de la página web, como su título, su contenido o su dirección, o incluso a los vínculos que dirigen a esa página.
- F) Dominios. El motor de búsqueda puede buscar páginas procedentes de cualquier servidor conectado a Internet incluido en su base de datos, pero si no nos interesa buscar en ciertos servidores o dominios, podemos o bien restringir la búsqueda a uno o varios dominios que sí nos interesan, o bien excluir uno o varios dominios que no nos interesan. De esa forma, acotamos la búsqueda por el criterio del dominio.

El último paso será la selección de resultados, ya que no todos los resultados ofrecidos por el motor de búsqueda serán útiles. Para saber si un resultado es útil, podemos pinchar en el enlace correspondiente y comprobar la temática, pero puede haber muchos enlaces que no nos aporten información de interés y puede llevarnos mucho tiempo entrar en cada una de esas páginas. Para agilizar la localización de los sitios más adecuados a nuestro objetivo es conveniente prestar un poco de atención a la información que se ofrece en el listado de resultados antes de pinchar en el enlace y acceder a la página. En la Figura 2 podemos ver una

muestra de página de resultados, que comentamos a continuación.

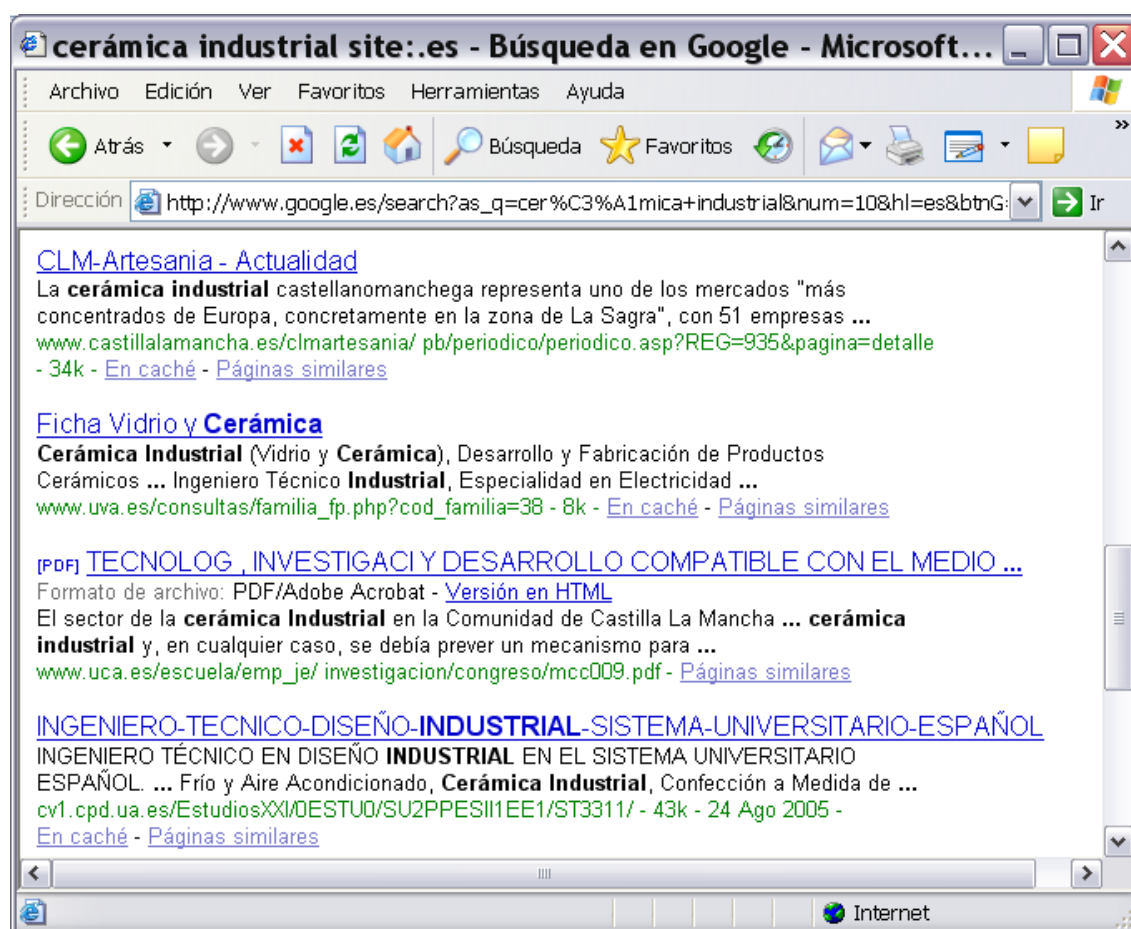


Figura 2. Ventana de resultados de búsqueda en Google

- A) Título. El título muestra una breve descripción del contenido de la página. Este título sólo se muestra en esta página del buscador porque está insertado en la página como una etiqueta invisible. En ocasiones, se muestra <No Title> porque el creador de la página no ha introducido esta etiqueta.
- B) Descripción. La descripción de una página puede haber sido implementada como una etiqueta en la propia página web por el desarrollador del sitio. Si no existe en la página como etiqueta, el buscador muestra las primeras líneas de texto que encuentra en esa página. Tanto el título como la descripción juegan un papel importante para decidir si un sitio resulta interesante o no para nuestros objetivos.
- C) URL. En ocasiones, el sitio web puede no estar accesible en un momento determinado, por lo que es interesante que en la página de resultados aparezca no sólo el hipervínculo, sino también la dirección completa URL, de manera que podamos apuntarla para visitar la página en otro momento.

- D) Fecha de actualización. La fecha de actualización de una página nos indica si el sitio se actualiza y por tanto contiene información y terminología actual. Puede ser interesante especialmente cuando se trata de informes anuales o leyes. No obstante, la fecha de actualización no indica necesariamente que la página se haya revisado en profundidad.
- E) Tamaño. La indicación del tamaño del fichero nos dará una idea del tiempo que tardaremos en descargar el documento.
- F) Relevancia del sitio. Algunos buscadores muestran, en porcentaje, el grado de afinidad del sitio con la petición realizada. Cuanto más se aproxime al 100%, más se ajusta el documento a nuestra búsqueda.
- G) Lengua. El buscador determina la lengua del sitio comparando las palabras que contiene con sus diccionarios, y en ocasiones puede haber errores.
- H) Grupo de sitios. Los buscadores pueden agrupar las páginas encontradas que están alojadas en el mismo sitio web, con lo que visitando uno de ellos ya podemos saber si el sitio nos interesa o no.

Si la página de resultados nos muestra demasiados resultados o muy pocos resultados, deberemos volver a plantear la pregunta, ajustando los criterios de búsqueda.

Si tenemos en cuenta todos estos aspectos al hacer las búsquedas de textos especializados podremos ahorrar tiempo en nuestras búsquedas y nuestro trabajo será más productivo y eficaz.

2.1.2. Recopilaciones de información seleccionada

En lugar de hacer búsquedas discriminadas sólo a través de palabras clave, podemos optar por aprovechar las selecciones y recopilaciones de fuentes elaboradas por investigadores u otros profesionales que pueden estar disponibles en forma impresa como artículos especializados o bien en Internet en forma de enlaces correctamente seleccionados y clasificados. Efectivamente, las dificultades que entraña la búsqueda y evaluación de fuentes fiables en Internet que sean útiles para el trabajo del traductor especializado ha llevado a investigadores y profesionales a publicar los resultados de estas indagaciones en forma de artículos en revistas especializadas en traducción o afines. Algunos manuales sobre traducción especializada o documentación en Traducción recogen buen número de este tipo de recopilaciones de gran utilidad. Como ejemplos podemos citar los manuales de Gonzalo García y García Yebra, 2004, Corpas Pastor, 2003, Gonzalo García y García Yebra, 2000, Maciá, 1998, Cordon García *et al.*, 2001. También las revistas especializadas y profesionales

y las actas de congresos recogen en ocasiones artículos con este tipo de información.

Por otra parte, disponemos también en Internet de recopilaciones de enlaces a sitios web de interés para el traductor, en forma de sección de enlaces de páginas personales de traductores o de empresas de traducción, portales de traducción (como Babel <<http://www.babelfish.com>>, Proz <<http://www.proz.com>> o Aquarius <<http://www.aquarius.net>>). Cabe destacar la recopilación hecha por Gonzalo (Gonzalo García, 2003) en DocuTradSo, propuesta de clasificación de fuentes de información en línea para la actividad traductora. En estos listados se recopilan enlaces de páginas relevantes por su temática y fiables por su autoría y vocación de permanencia y actualización periódica. La publicación de estas recopilaciones en Internet facilita su actualización periódica, lo que resulta muy útil ya que muchas páginas pueden cambiar de dirección y obviamente estos cambios no quedan reflejados en las recopilaciones publicadas de forma impresa.

Una forma similar de acceso a la información previamente seleccionada es consultar los temarios de asignaturas que publican en Internet las universidades. Estos programas suelen contener una sección de bibliografía e incluso indicación de sitios web de interés para la materia. Constituyen por lo tanto fuentes de información relevante y actual sobre la materia de que se trate y por tanto resultarán plenamente fiables.

2.1.3. Acceso a expertos

Otra de las formas seguras de acceder a información fiable es que sean los mismos expertos en un ámbito temático los que nos proporcionen las referencias más relevantes en ese ámbito temático, ya sean fuentes impresas o digitales. El correo electrónico, las listas de distribución y boletines de noticias, y los centros de expertos constituyen los principales servicios que pone Internet a nuestra disposición para contactar directamente con expertos. Las diferencias entre el uso de uno de estos servicios es el tipo de relación que se establece entre el emisor y receptor de la pregunta, especialmente en cuanto al compromiso de ofrecer la ayuda solicitada.

A) Correo electrónico. Mediante el correo electrónico, el traductor puede contactar directamente con el experto porque ha encontrado sus datos a través de un directorio de expertos, su página web personal, la página web de una universidad o empresa, etc. Como se trata de un contacto personal, no existe ningún compromiso del experto en contestar a la información solicitada, por lo que la posibilidad de que el experto responda o no, así como el grado de satisfacción de la respuesta que ofrezca dependerá de distintos parámetros. No

obstante, la experiencia de estudiantes de traducción que se ponen en contacto con expertos es bastante positiva. Para garantizar la máxima adecuación en la respuesta, la pregunta debe reunir algunos requisitos como presentar bien el contexto o las razones por las que nos dirigimos a ese experto, señalar las búsquedas que se han hecho con anterioridad, formular correctamente la pregunta y, por supuesto, ser educado y amable en la expresión, teniendo en cuenta que nos dedica su tiempo de manera altruista.

- B) Listas de distribución y boletines de noticias. En el caso de las listas de distribución y boletines de noticias, los participantes han decidido voluntariamente suscribirse a mensajes relacionados con un determinado ámbito temático. En algunas de ellas es necesario demostrar el perfil profesional o la razón profesional que lleva al solicitante a participar en esa lista. Además, pueden estar moderadas. Cualquier pregunta lanzada a la lista será vista por todos los suscritos, quienes podrán responder al mensaje que a su vez también será visto por el resto de suscritos y a su vez podrán completar la información. Todo ello garantiza en cierta medida la fiabilidad de la información obtenida. Existen listas de distribución nacionales o internacionales en las que los traductores intercambian información sobre diferentes aspectos de su profesión, incluyendo aspectos de información terminológica y documentación. Probablemente la más activa es Traducción en España, pero hay también muchas otras (Alcina Caudet, 2003). Si lo que se desea es entrar en listas de distribución sobre ámbitos temáticos diferentes de la traducción (por ejemplo, un traductor que desee hacer consultas sobre el ámbito de la medicina) lo más adecuado será consultar al moderador de la lista si es posible suscribirse para hacer este tipo de consultas, ya que lo normal es que la comunicación en las listas sea de experto a experto.
- C) Centros de expertos. Esta puede ser una buena solución para el tipo de consultas que suelen necesitar los traductores, ya que en este tipo de servicio la comunicación es entre experto y lego, pero existe el compromiso por parte del experto para dar una respuesta satisfactoria a la persona que formula la pregunta. La situación ideal es la de centros de expertos creados específicamente para traductores, como es el caso de Guía de Expertos de la UMA para traductores (<<http://www.ieev.uma.es/areadoc/EXPERTOS/index.html> >) creada por Palomares, ya que de esta manera el experto conoce de antemano el perfil del usuario, es decir, las razones y las principales problemas con que se suele enfrentar el traductor, y por tanto puede ajustar mejor las respuestas. No obstante, existen otros centros de expertos que también pueden servir a los fines del traductor (por ejemplo, centros de expertos creados por asociaciones que desean llevar a cabo una labor divulgativa de sus fines como el mejor conocimiento de una enfermedad, de un deporte o de una actividad

económica).

2.2. Evaluación de los resultados

En general, lo más importante para el traductor es acceder a páginas con terminología propia y precisa de una especialidad. Necesita encontrar o comprobar en esas páginas el significado de un término, su uso en contexto o su equivalencia a otra lengua. En el caso de las fuentes impresas, sabemos que los libros y artículos científicos que se publican han pasado el filtro de un comité editorial que estudia y acepta la propuesta de un autor de publicar un libro o un artículo científico, un editor que revisa el texto, y un proceso editorial que asegura la correcta redacción. Por otra parte, también hemos aprendido a diferenciar que las editoriales (o sus colecciones) pueden especializarse en un tema, o estar dirigidas a un determinado público (libros científicos, libros de texto, libros divulgativos), y todos estos criterios se tienen en cuenta en la elaboración de los corpus especializados (Cabré, 1992; Auger y Picht, 1987; Adelstein, 2002). Estos parámetros que utilizamos para seleccionar o clasificar fuentes impresas debemos aplicarlos también cuando se trata de páginas que encontramos en Internet. En el caso de las páginas de Internet es el usuario –en nuestro caso el traductor o terminólogo– quien debe habituarse a valorar y evaluar los sitios y las páginas encontradas.

Al igual que en el caso de las fuentes impresas, hay varios parámetros que nos ayudan a comprobar la fiabilidad de las fuentes electrónicas. El grado de fiabilidad de una página resultará de la valoración de distintos parámetros, como autoría, lengua original del texto, actualidad y servidor en que se aloja.

A) Autor. El autor debe ser una fuente autorizada y experta en la materia (especialista en el tema, universidad). El hecho de que otros expertos recomienden esa página es un dato que por sí mismo otorgará fiabilidad a la página. Una manera de comprobar si la página encontrada es recomendada por otros expertos es pidiendo al motor de búsqueda los hipervínculos que dirigen a esa página. Por ejemplo, si queremos comprobar cuántos hipervínculos dirigen a la página de la Asociación Española de Cerámica y Vidrio <<http://www.spaintiles.info>>, simplemente escribimos en la casilla de google lo siguiente link: <<http://www.spaintiles.info>>, y nos devolverá más de 30 resultados de sitios web que contienen ese enlace. Esto nos indica por una parte la estabilidad del sitio (si no llevara un cierto tiempo en Internet no existirían enlaces a esa página), y también su especialización y fiabilidad (podemos comprobar que las páginas que enlazan con esta página son también

especializadas en cerámica y están respaldadas también por instituciones y expertos).

- B) Lengua original. Desde el punto de vista del buen uso de la terminología, es importante constatar que el texto es original y no se trata de una traducción, ya que al traducir se pueden generar giros y expresiones que, no siendo incorrectos, no responden al uso que haría un experto de esa terminología. Además, pocas veces se puede comprobar quién es el autor de la traducción.
- C) Actualidad. En Internet, existen páginas que pueden haber sido creadas y publicadas y no haber sido actualizadas jamás, por lo que pueden tener varios años de antigüedad. Esta circunstancia puede indicarnos una menor fiabilidad de la información que contienen, ya que Internet se concibe como un medio de publicación dinámico y se espera que los contenidos se cuiden y se mantengan actualizados. La fecha de actualización puede aparecer indicada en algún lugar de la página, pero no siempre es así.
- D) El servidor que aloja la página. La dirección de Internet en que aparece la página nos indica la empresa o institución que está dando el apoyo necesario para que esa página esté publicada en Internet. Por ejemplo, la dirección https://e-ujier.uji.es/pls/www/!gri_ass.leu050003?pp_activa=&p_desde=&id_asignatura=IS05&id_titulacion=28&activa=9&lg=ES presenta el programa docente de la asignatura *Fundamentos matemáticos de la informática* que se imparte en la Universitat Jaume I. El hecho de que sea una institución universitaria la que aloja esta página le confiere mayor fiabilidad a los contenidos porque obviamente éstos habrán sido cumplimentados por el profesorado de la propia institución. Podemos comprobarlo porque la dirección “uji.es” nos indica que el sitio corresponde a la Universitat Jaume I. En cambio, si encontramos una página alojada en un sitio como “yahoo.com” no podremos otorgarle la misma fiabilidad, ya que la empresa Yahoo ofrece el servicio de alojar páginas web a cualquier persona que lo solicite y no comprueba (no es su función hacerlo, salvo casos de delito) ni la identidad de las personas ni la información que ofrecen en sus páginas. Es decir, que si la página está firmada por una persona que dice ser médico, esto podrá ser cierto, pero realmente nosotros no tenemos ningún medio para verificarlo, por lo que el grado de fiabilidad que podemos otorgar a la información que allí aparece es bajo.

En la ficha de evaluación de textos especializados que hemos utilizado en el proyecto CREC y que adjuntamos en un anexo a este capítulo, hemos intentado recoger todos los aspectos relacionados con la identificación de la fuente (impresa o electrónica), la autoría y grado de fiabilidad del autor y del sitio, en los apartados:

- Datos de la referencia bibliográfica: datos necesarios para la identificación de la fuente ya sea impresa o electrónica.
- Datos generales: hacen referencia a la lengua en que está escrito el texto, si existen textos paralelos y si está disponible en formato electrónico.
- Fiabilidad y calidad de la referencia: se indagan cuestiones relacionadas con la fiabilidad del autor y del sitio.
- Aspectos pragmáticos y grado de especialización del texto: se indagan cuestiones relacionadas con el nivel de especialización del texto, el emisor y el receptor.
- Aspectos lingüísticos: se indagan las características sintácticas, morfológicas y estilísticas del texto para asegurarnos de su pertenencia a un lenguaje de especialidad y de un grado de especialización.

Cada uno de estos apartados contienen diferentes aspectos a analizar. Lo que se pretende a través de esta ficha es acostumbrar al estudiante de traducción a analizar los distintos aspectos de las fuentes documentales en relación con la información que interesa desde el punto de vista de la terminología y la traducción.

2.3. Almacenamiento de la información

Una vez hemos realizado la búsqueda y evaluación, para no perder el trabajo realizado podemos tomar varias opciones: 1) guardar la consulta que hemos hecho al buscador en una carpeta especial de nuestros favoritos, de manera que si tenemos que interrumpir la fase de evaluación de los resultados podamos continuarla en otro momento; 2) una vez visitado un sitio y comprobada su utilidad e interés para nuestro objetivo, podemos guardar esa dirección en una carpeta especial de nuestros enlaces favoritos con el nombre del proyecto; 3) complementariamente a guardar la dirección de la página en favoritos, también podemos guardar la página (es decir, el fichero htm/html) en una carpeta especial de nuestro disco duro, con el nombre del proyecto. En ocasiones, puede que nos interese no sólo guardar una página concreta, sino el conjunto de páginas comprendidas en un sitio web. En este caso, existen programas específicos (muchos de ellos gratuitos) que nos ayudarán a descargar automáticamente un conjunto de ficheros, sin tener que guardar una a una cada página interesante (un ejemplo de este tipo de programas es Htrack).

El almacenamiento de ficheros en nuestro disco duro es un procedimiento cada vez más recomendable. El motivo principal es que tener los textos en nuestro propio ordenador nos permitirá procesarlos con programas específicos de análisis textual y concordancias

(como WordSmith) o de extracción terminológica (como Term Extract de Trados) que nos permiten realizar distintos tipos de búsquedas terminológicas y fraseología, colocaciones, obtención de estadísticas de frecuencia, etc. Todo ello, en un corpus textual creado específicamente para cumplir nuestros objetivos (Sánchez-Gijón, 2005). En cambio, un programa como InfoRapid permite obtener concordancias directamente de cualquier sitio de Internet, sin necesidad de buscar mediante motores de búsqueda y almacenar los documentos en nuestro ordenador, pero eso implica también que no nos aseguramos la fiabilidad de las fuentes de esas concordancias.

Además de obtener textos en formato electrónico a través de Internet, puede ser necesario complementar el corpus textual con documentos impresos relevantes por su temática y fiabilidad que habrá que *digitalizar*, es decir, escanear y convertir a formato de texto.

Generalmente, antes de ser procesados a través de un programa de análisis textual o de extracción terminológica, habrá que homogeneizar el formato de los documentos de texto (por ejemplo, al formato txt o rtf, según el formato admitido por esos programas). Para ello, deberemos disponer también de programas convertidores de formato que permitan eliminar las etiquetas internas propias del lenguaje HTML de los documentos HTML o convertir un fichero PDF a texto crudo como TXT o texto enriquecido (por ejemplo, PDFtoTXT, u otros).

3. Hacia el futuro en la documentación del traductor

La diversificación constante de disciplinas y especialidades y el aumento considerable de conocimiento que se produce en cada una de ellas hacen que la búsqueda, selección y almacenamiento de textos especializados sea una labor cada vez más ardua para el traductor especializado, a pesar de las aparentes ventajas que ofrece hoy en día Internet. El traductor puede llevar a cabo estas tareas de forma limitada. Debido a la imposibilidad económica de suscribirse individualmente a muchos de los servicios de información y bases de datos relevantes para su labor (en las que la información ya está evaluada, seleccionada y clasificada), muchos documentos relevantes son inaccesibles. De cara al futuro, lo interesante para el traductor o mediador lingüístico, será que todo este proceso de información se desarrolle a gran escala y se ofrezca en Internet para los profesionales interesados. Habrá que aunar los esfuerzos tanto de traductores profesionales y empresas de traducción en la recopilación de esta información.

Pensamos que una buena estrategia será la creación de plataformas en Internet desde

las que se ofrezca y se obtenga la información. Estas plataformas podrían ser diseñadas y mantenidas por traductores y empresas de traducción bajo la forma similar a entornos de trabajo colaborativo (Alcina Caudet, 2002), o bien podrían ser diseñadas y mantenidas por empresas de información que desarrollen servicios específicos de apoyo a los traductores mediante suscripción u otras fórmulas, al igual que se han desarrollado otros servicios de información para otros profesionales, como las bases de datos de legislación para abogados y juristas, o las bases de datos bibliográficas para los investigadores.

El desarrollo de estos sistemas de apoyo a la labor del traductor podrían desarrollarse de diferentes formas, ya sea como Directorios o Portales de traducción especializada, o bien como Motores de búsqueda centrados en las necesidades específicas del traductor especializado (de similar modo está concebido *Google scholar*, dirigido especialmente al colectivo de investigadores). Probablemente en un futuro próximo las tecnologías GRID (Global Resources Information Database) podrán contribuir proporcionando poderosas herramientas. En este tipo de plataformas se ofrecería acceso a información relevante para el traductor, como 1) textos especializados recopilados previamente de forma manual o automática y clasificados según el ámbito temático, el tipo de texto, el género textual, la lengua; 2) diccionarios y bancos de datos terminológicos. No sólo se trataría de ofrecer los textos especializados clasificados, sino también concordancias de términos, textos paralelos, indicación de niveles de especialidad del uso de un término o modelos de textos de un determinado género textual en una determinada lengua.

Para la consecución de estos objetivos, será necesaria la colaboración de traductores y empresas de traducción que expliciten sus necesidades de documentación e información, expertos en lingüística computacional que desarrollen sistemas de recuperación de información lingüística y traductológica en línea y expertos en información y documentación e informáticos que contribuyan a la mejor clasificación de los textos y fragmentos de información.

BIBLIOGRAFÍA

- Adelstein, A. (2002) "Estatuto lingüístico del término", en J. García Palacios y M.T. Fuentes Morán (eds.) *Texto, Terminología y Traducción*. Salamanca: Ediciones Almar.
- Alcina Caudet, M.A. (2002) Tutorizar trabajos terminológicos en un entorno virtual de trabajo colaborativo. *Sendebarr* 13, 169-181.

- Alcina Caudet, M.A. (2003) Encouraging the use of e-mail and mailing lists among translation students. *Meta* 48 (4), 634-641.
- Auger, P. y Picht, H. (1987) *Metodologia de la recerca terminològica*. Barcelona: Generalitat de Catalunya.
- Bergeron, M. y van Steenberghe, M. (2000) The Internet Search Tutorial (for Translators). Online documents at <http://www.mabcom.com/websearch/tutorial.html>. [01-07-2005].
- Budin, G. y Wright, S.E. (2001) "Terminology Management Resources on the Internet", en S.E. Wright y G. Budin (eds.) *Handbook of Terminology Management*. Philadelphia: John Benjamins, 845-858.
- Cabré, M.T. (1992) *La terminologia. La teoria, els mètodes, les aplicacions*. Barcelona: Empúries.
- Cordón García, J.A.; López Lucas, J. y Vaquero Pulido, J.R. (2001) *Manual de investigación bibliográfica y documental. Teoría y práctica*. Madrid: Pirámide.
- Corpas Pastor, G. (ed.) (2003) *Recursos documentales y tecnológicos para la traducción del discurso jurídico (español, alemán, inglés, italiano, árabe)*. Granada: Comares.
- Gonzalo García, C. y García Yebra, V. (eds.) (2000) *Documentación, terminología y traducción: Síntesis*.
- Gonzalo García, C. (2003) DocuTradSo. Online documents at <http://www3.uva.es/DocuTradSo/>.
- Gonzalo García, C. y García Yebra, V. (eds.) (2004) *Manual de documentación y terminología para la traducción especializada*. Madrid: Arco Libros.
- Maciá, M. (ed.) (1998) *Manual de documentación jurídica*. Madrid: Síntesis.
- Sánchez-Gijón, P. (2005) *L'ús de corpus en la traducció especialitzada: Compilació de corpus ad hoc o extracció de recursos terminològics*. Girona: Documenta Universitaria.

Preprint del capítulo de libro:

Alcina, A., V. Soler y A. Estellés (2005): «Internet como instrumento para la documentación en terminología y traducción. Hacia las plataformas de recursos electrónicos para el traductor especializado», en D. Sales (ed.) *La Biblioteca de Babel: Documentarse para traducir*, Granada: Comares, 221-241. ISBN: 84-8151-032-7. <http://hdl.handle.net/10234/192441>
@ Las autoras

ANEXO I

FICHA DE EVALUACIÓN PROYECTO CREC

3.1.1.1. datos de la referencia bibliográfica

Código:

Estudiante:

Tema: ☐ Copa América ☐ Cerámica

Rama (ver árbol de campo):

¿La obra está accesible? ☐ Sí ☐ No

Ubicación del ejemplar: ☐ Biblioteca UJI
☐ Libro propio
☐ Instituto de Cerámica
☐ Otra. ¿Cuál?

¿Está disponible en formato electrónico? ☐ Sí ☐ No

Si está disponible en formato electrónico indicar dónde:
.....

Libro	Título: Autor/es: Año de primera publicación: Editorial: Lugar de publicación: Número de edición y año:
Capítulo de libro	Título del capítulo: Autor/es del capítulo: Páginas: Título del libro: Editor/es del libro: Editorial: Año de publicación: Lugar de publicación: Número de edición y año:
Artículo de revista	Título del artículo: Autor/es: Título de la revista: Lugar de publicación: Volumen: Número: Año de publicación: Páginas:
Página Web	Título de la página: Autor/es: Hipervínculo al sitio Web: Fecha de consulta:

DATOS GENERALES

1. ¿En qué lengua está escrita? ☐ Español ☐ Catalán ☐ Inglés
2. ¿Se trata de una traducción o del original? ☐ Original ☐ Traducción
 - a) En caso de ser una traducción, ¿está disponible el original? ☐ Sí ☐ No
¿Dónde?
 - b) En caso de ser el texto original, ¿has localizado alguna traducción? ☐ Sí ☐ No
¿A qué lengua?.....
¿Dónde?
 - c) ¿Está disponible en formato electrónico el texto paralelo? ☐ Sí ☐ No
¿Dónde?

FIABILIDAD Y CALIDAD DE LA REFERENCIA

3. ¿Dónde/Cómo has encontrado la referencia?
 - ☐ Bibliografía del programa de una asignatura
¿Qué asignatura?.....
 - ☐ Catálogo de una biblioteca universitaria
¿Qué universidad?
 - ☐ Página Web de una asociación, organización, etc.
¿Qué asociación, organización, etc.....
 - ☐ Recomendada por un especialista
Datos del especialista.....
 - ☐ Página Web personal. ¿Cuál es el nombre, el cargo y el lugar de trabajo?
.....
 - ☐ Otros. ¿Cuáles?

ASPECTOS PRAGMÁTICOS Y GRADO DE ESPECIALIZACIÓN DEL TEXTO

4. ¿Quién es el autor?
 - ☐ Es experto. ¿Por qué?.....
 - ☐ Es aficionado. ¿Por qué?
 - ☐ Otros. ¿Quién?
5. Sobre la publicación, se trata de...
 - ☐ Una editorial de prestigio en este ámbito
 - ☐ Una editorial que pertenece a una universidad
 - ☐ Una editorial con amplio prestigio
 - ☐ Un editor con reconocido prestigio
 - ☐ Otros. ¿Cuáles?
6. ¿Quiénes son los destinatarios de la obra?
 - ☐ Profesionales /expertos
 - ☐ Público en general
 - ☐ Estudiantes de esta materia
 - ☐ Otros. ¿Cuáles?
7. ¿Qué rasgos te permiten reconocer el público al que va dirigida la obra?
 - ☐ Se indica en el prólogo
 - ☐ Por el tipo de publicaciones de la Editorial
 - ☐ Otros. ¿Cuáles?

ASPECTOS LINGÜÍSTICOS DEL TEXTO

CARACTERÍSTICAS SINTÁCTICAS DEL TEXTO

Marca con una cruz si el texto reúne alguna de estas características e indica algún ejemplo:

- ☐ Frases más cortas y poca subordinación compleja.
.....
- ☐ Uso de la primera persona del plural como fórmula de modestia
.....
- ☐ Ausencia de exclamaciones
.....
- ☐ Uso de frases cortas
.....
- ☐ Falta de redundancias innecesarias
.....
- ☐ Uso frecuente de fórmulas impersonales
.....
- ☐ Atención centrada en los sintagmas nominales
.....
- ☐ Referencias a otros sistemas de representación incorporados en el cuerpo del texto (gráficos, fórmulas, nomenclaturas...).
- ☐ Otras. ¿Cuáles?

CARACTERÍSTICAS MORFOLÓGICAS DEL TEXTO

Marca con una cruz si el texto reúne alguna de estas características e indica algún ejemplo:

- ☐ Estructuras morfológicas compuestas de formantes cultos. Ejemplos:
- ☐ Sinapsis (sintagmas terminológicos). Ejemplos:
- ☐ Siglas, símbolos y neologismos especializados. Ejemplos:
- ☐ Nominalizaciones a partir de verbos. Ejemplos:
- ☐ Otras características. ¿Cuáles?

CARACTERÍSTICAS ESTILÍSTICAS DEL TEXTO

Marca con una cruz si el texto reúne alguna de estas características e indica algún ejemplo:

Relacionados con la precisión

- ☐ Utilización de terminología propia de la especialidad
.....

Relacionados con la claridad

- ☐ Predominio de la coordinación y yuxtaposición sobre la subordinación.
.....
- ☐ Oraciones subordinadas adjetivas explicativas
.....
- ☐ Uso de incisos entre comas, rayas y paréntesis:
.....
- ☐ Aposiciones
.....
- ☐ Uso de la conjunción o con el valor de identificación o equivalencia:
.....
- ☐ Definiciones, aclaraciones o resúmenes tras dos puntos:
.....
- ☐ Enlaces explicativos (es decir, esto es, a saber, etc.) que consolidan la coherencia textual.
.....
- ☐ Repetición de palabras
.....
- ☐ Elementos ordenadores del pensamiento que distribuyen las acciones en secuencias
.....
- ☐ Elementos ordenadores del pensamiento que establecen relaciones de consecuencia o conclusión.
.....

Relacionados con la objetividad

1. Diluyendo la importancia del autor

- ☐ Oraciones enunciativas en función referencial
.....
- ☐ Oraciones pronominales con se con valor impersonal o de pasiva refleja.
.....
- ☐ Oraciones de pasiva analítica (con ser).
.....
- ☐ Nominalización de frases verbales o sustitución de la construcción verbal por la nominal
.....
- ☐ Nominalización del infinitivo.
.....

2. Destacando los hechos y los datos

- ☐ Utilización de adjetivos especificativos (generalmente pospuestos, en especial los adjetivos utilizados para clasificar al sustantivo).
.....
- ☐ Utilización de adjetivos especificativos que se han lexicalizado en un sintagma:
.....
- ☐ Oraciones de relativo especificativo
.....
- ☐ Complementos preposicionales del nombre

-
- ☐ Uso predominante del indicativo.
-
- 3. Determinando las circunstancias que acompañan a los procesos
- ☐ Distintos tipos de oraciones adverbiales.
-
- ☐ Los complementos circunstanciales
-
- ☐ Infinitivos, gerundios y participios, al comenzar la frase, expresando circunstancias temporales:
-
- ☐ Gerundio, al comenzar la frase, con valor condicional o consecutivo y modal
-
- ☐ Participio, al comenzar la frase, con valor condicional o consecutivo
-
- 4. Relacionados con la universalidad
- ☐ Uso del artículo con valor generalizador.
-
- ☐ Uso del presente atemporal
-
- ☐ Presencia de nombres abstractos
-
- ☐ Uso de tecnicismos
-